



# THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### The Experimental Design Assistant

**Citation for published version:**

Percie du Sert, N, Bamsey, I, Bate, ST, Berdoy, M, Clark, RA, Cuthill, I, Fry, D, Karp, NA, Macleod, M, Moon, L, Stanford, SC & Lings, B 2017, 'The Experimental Design Assistant', *PLoS Biology*, vol. 15, no. 9, e2003779. <https://doi.org/10.1371/journal.pbio.2003779>

**Digital Object Identifier (DOI):**

[10.1371/journal.pbio.2003779](https://doi.org/10.1371/journal.pbio.2003779)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

PLoS Biology

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



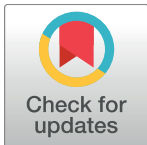
COMMUNITY PAGE

# The Experimental Design Assistant

**Nathalie Percie du Sert<sup>1\*</sup>, Ian Bamsey<sup>2</sup>, Simon T. Bate<sup>3</sup>, Manuel Berdoy<sup>4</sup>, Robin A. Clark<sup>5</sup>, Innes Cuthill<sup>6</sup>, Derek Fry<sup>7</sup>, Natasha A. Karp<sup>8,9</sup>, Malcolm Macleod<sup>10</sup>, Lawrence Moon<sup>11</sup>, S. Clare Stanford<sup>12</sup>, Brian Lings<sup>2</sup>**

**1** National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs), London, United Kingdom, **2** Certus Technology Associates Ltd, Exeter, United Kingdom, **3** GlaxoSmithKline, Stevenage, United Kingdom, **4** University of Oxford, Oxford, United Kingdom, **5** Envigo, Alconbury, United Kingdom, **6** University of Bristol, Bristol, United Kingdom, **7** University of Manchester, Manchester, United Kingdom, **8** Wellcome Trust Sanger Institute, Hinxton, United Kingdom, **9** Quantitative Biology IMED, AstraZeneca R&D, Cambridge, United Kingdom, **10** Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, United Kingdom, **11** King's College London, London, United Kingdom, **12** University College London, London, United Kingdom

\* [nathalie.perciedusert@nc3rs.org.uk](mailto:nathalie.perciedusert@nc3rs.org.uk)



## Abstract

Addressing the common problems that researchers encounter when designing and analysing animal experiments will improve the reliability of in vivo research. In this article, the Experimental Design Assistant (EDA) is introduced. The EDA is a web-based tool that guides the in vivo researcher through the experimental design and analysis process, providing automated feedback on the proposed design and generating a graphical summary that aids communication with colleagues, funders, regulatory authorities, and the wider scientific community. It will have an important role in addressing causes of irreproducibility.

## OPEN ACCESS

**Citation:** Percie du Sert N, Bamsey I, Bate ST, Berdoy M, Clark RA, Cuthill I, et al. (2017) The Experimental Design Assistant. *PLoS Biol* 15(9): e2003779. <https://doi.org/10.1371/journal.pbio.2003779>

**Published:** September 28, 2017

**Copyright:** © 2017 du Sert et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The authors received no specific funding for this work.

**Competing interests:** The EDA was developed as part of an NC3Rs programme. All authors were involved in developing the EDA and NPdS currently works at the NC3Rs.

**Abbreviations:** EDA, Experimental Design Assistant; NC3Rs, UK National Centre for the Replacement, Refinement and Reduction of Animals in Research; ARRIVE, Animal research: Reporting of in vivo experiments.

**Provenance:** Commissioned; externally peer reviewed.

## Introduction

The poor reproducibility of findings from animal research has received much attention over the last few years [1], not least because of the impact it has on translation, scientific progress, and the use of resources. It has been estimated that over half of preclinical research is irreproducible (see [2]). There are many reasons for this, aside from the complication that reproducibility can be defined in different ways [3], but flawed experimental design, inappropriate statistical analysis, and inadequate reporting have been flagged as major concerns [4–6]. There is considerable scope for improving the way animal research is designed, conducted, analysed, and reported.

As a starting point, the UK National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs) developed the Animal research: Reporting of in vivo experiments (ARRIVE) guidelines to improve the reporting of animal experiments [7,8]. As of 2016, compliance with these guidelines was recommended by over 600 leading journals in the biomedical sciences [9]; with more advocating their use each year, this number has now increased to over 1,000 [7]. Compliance should ensure that published articles contain sufficient information to assess the reliability of the findings and enable the experiments to be adequately replicated [8]. Improved reporting should also increase the quality of retrospective studies,

such as systematic reviews. However, in order to increase the reliability of findings, the design, conduct, and analysis of individual experiments needs to be improved. Here, we present the Experimental Design Assistant (EDA), which was launched to support the scientific community with this process [10,11].

The EDA (<https://eda.nc3rs.org.uk>) is a web-based application with an integrated website, which guides researchers through the process of designing animal experiments; the output includes a diagram that improves the transparency of the experimental plan. The resource is freely available and was developed by the authors as an NC3Rs-led collaboration between in vivo researchers and statisticians from academia and industry and a team of software designers who specialise in innovative solutions for the life sciences (<http://www.certus-tech.com/>). The EDA enables researchers to build a stepwise, schematic representation of an experiment—the EDA diagram—and uses computer-based logical reasoning to provide feedback and advice on the experimental plan. The system's main features are presented in Table 1.

## The EDA improves the reliability of experimental results and analysis

The majority of published in vivo research provides no indication that basic precautions have been taken to obtain reliable findings [4,12]. High internal validity can only be achieved by minimising systematic bias so that observed differences can be confidently associated with the treatment of interest. For example, many publications include no information on randomisation and blinding [4,12]. This is not considered to be just a reporting issue; studies have shown that publications of animal experiments that report the use of such measures also tend to report lower effect sizes compared to those that do not [13]. This implies that experiments are generally not designed and conducted to the highest standards, and the results may not be reliable. Random allocation and blinding have 2 benefits: first, they help meet a key assumption of the statistical analysis, namely, that different groups are drawn from the same background population using random sampling. Second, applying these techniques minimises systematic differences between the treatment groups during the conduct of the experiment, assessment of the results, and data analysis. Such differences can be caused by researchers subconsciously influencing the animals' allocation to treatment groups, the animals' behaviour [14], or the handling of the data (e.g., removal of outliers).

Another common concern regarding the reliability of animal experiments is that they are underpowered, using too few animals to yield dependable results. Button, Ioannidis and colleagues [15] estimated the average power in neuroscience animal studies to be around 20%. This constitutes a high risk of missing a genuine effect (a false negative), because only 1 in 5 experiments would have a chance of detecting an effect of the magnitude reported. Conversely, the use of sample sizes that are too small also reduces the reliability of the conclusions from an individual experiment and of the published literature as a whole. When a statistically significant effect is detected, it is less likely to be genuine and its magnitude more likely to be overestimated

**Table 1. Features of the Experimental Design Assistant (EDA).**

**Features of the EDA include the following:**

- A computer-aided design tool to develop a diagram representing the experimental plan,
- feedback from an expert system on the experimental plan (the Critique),
- Analysis Suggestion,
- sample size calculation,
- randomisation sequence generation,
- support for allocation concealment and blinding,
- web-based resources to improve knowledge of experimental design and analysis.

<https://doi.org/10.1371/journal.pbio.2003779.t001>

[15,16]. Indeed, justification for the choice of sample sizes is rarely included in published papers [4,12].

The EDA helps avoid such pitfalls when designing *in vivo* experiments and improves the reliability of the results—and ultimately, their reproducibility. The system generates a randomisation sequence for the experiment, which takes into account any blocking factors included in the design and provides dedicated functionalities, such as support for blinding and sample size calculation, to assist researchers in following best practice (see Fig 1). A tailored critique provides suggestions on optimising the experimental plan. For example, it helps researchers to identify variables that could confound the outcome and provides advice on how to include them in the randomisation and the statistical analysis. Finally, once the researcher has addressed the feedback and is satisfied with the design, the system advises on which methods of statistical analysis are most appropriate. Designing experiments with the EDA encourages researchers to consider the sources of bias at the design stages of the experiment before the data are collected, ensuring a rigorous design that is more likely to yield robust findings that can be reproduced. The EDA can also be used as a teaching resource, thereby promoting a better understanding of the principles of experimental design at an early stage of the research training process. The process of building an EDA diagram familiarises students with the different components of the design and how they are connected. The visual representation of abstract concepts, such as the experimental unit, the independent variables, or data transformations, brings clarity that enables a detailed discussion of the experimental plan.

## The EDA offers a new standard notation for describing experiments

It is difficult to find a technical discipline that has not adopted a schematic, diagrammatic, or symbolic notation to improve communications and the recording of methodological detail. However, there are no universally accepted standards to describe the different components of an experimental design. Different terms can be used to describe the same things; for example, the outcome measure is also known as the dependent variable, the response variable, the outcome variable, or the variable of interest, which can easily be confused with the independent variable of interest (also known as the factor of interest or the predictor of interest). By contrast, the same terms can be used to describe different settings; for example, a ‘repeated measure design’ can imply a situation in which animals receive multiple treatments in a different order (sometimes described as a crossover design). However, it could also refer to a situation in which the response to a given treatment in each animal is measured over time or to a situation in which multiple responses are measured for each animal [17]. The EDA resolves this problem by helping the user generate unambiguous representations of these different designs using EDA diagrams (see Fig 2 and S1 Fig) and hence does not require knowledge or understanding of labels such as ‘repeated measure design’.

A central feature of the EDA is the development of standards for communicating experimental design. This has required the careful definition of the concepts used in experimental design together with an associated vocabulary of preferred terms. This was developed using an iterative approach and tested using a wide range of experimental plans from the published literature. The result is an ontology that supports the capture of every element of an experimental plan, from high-level information such as the hypotheses, effect of interest, and animal characteristics, to the practical steps carried out in the laboratory, as well as details on the variables included in the design and statistical analysis. This ontology has underpinned the development of a computer-aided design tool to support the experimental design process. The tool helps users develop EDA diagrams consistent with the ontology. The diagrams are unambiguous and more explicit than the text descriptions normally included in grant applications or journal



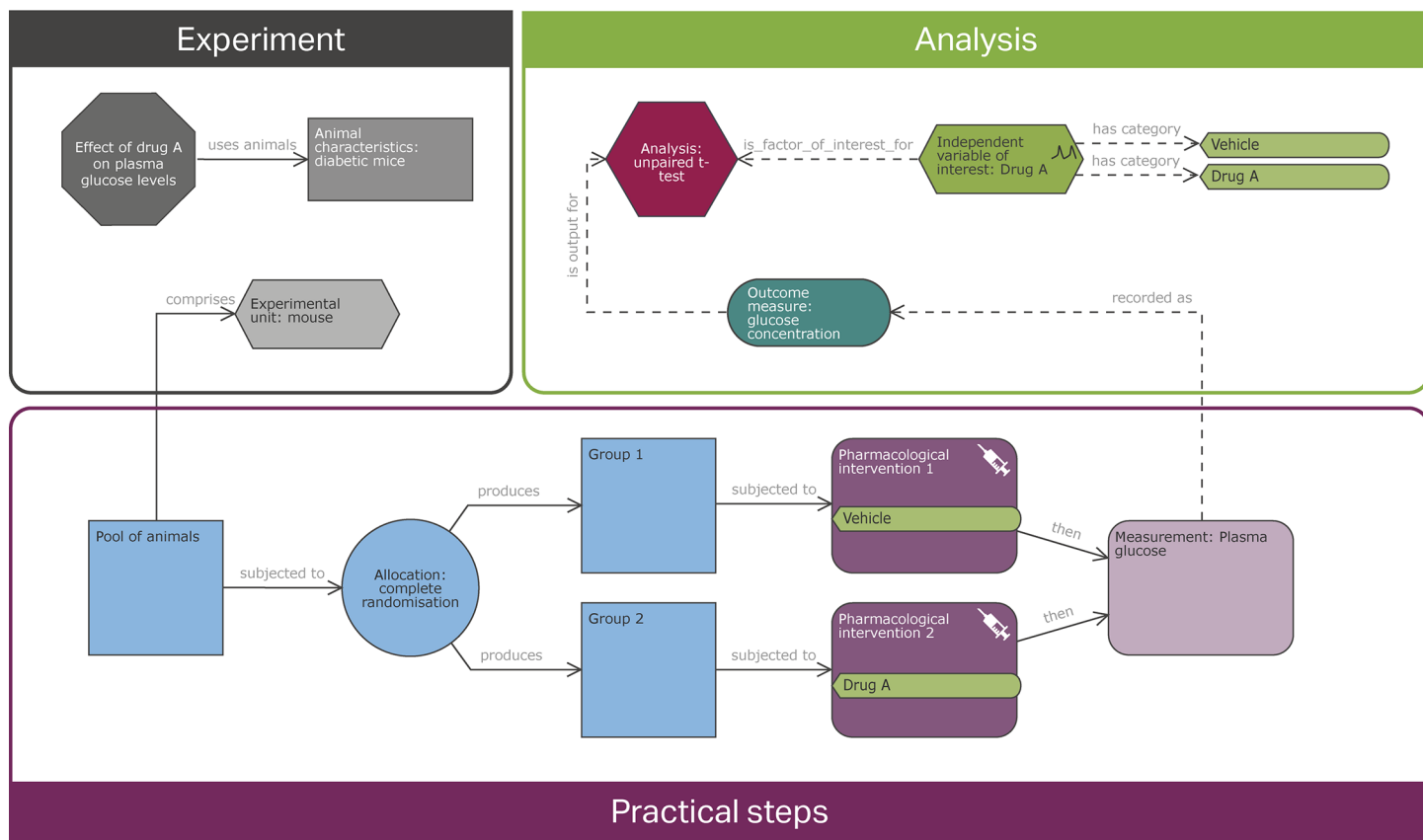
**Fig 1. The Experimental Design Assistant (EDA) workflow.** The workflow is not fixed and different users might prefer to do some steps in a different order. A potential workflow using the different functionalities of the EDA is described as follows: (1) The user starts by drawing a diagram (with nodes and links) representing the experiment they are planning. Assistance is provided in the form of examples, templates, and video tutorials. (2) Information is added into the node properties, providing more details about the specific step of the process represented by each node. (3) The “Critique” functionality (see Table 2) enables the researcher to obtain feedback on the diagram and the design it represents. The feedback might prompt a change in their plans or the addition of missing information. This is an iterative process and the user might go through the first 3 steps a number of times. (4) Once feedback from the critique has been addressed and the user is satisfied with the final design, the analysis method suggested by the system can be reviewed (see Table 2). (5) Depending on how the data will be analysed, a suitable sample size can be calculated using one of the calculators provided within the system. (6) Once the number of animals needed per group is known, the EDA can generate the randomisation sequence. The spreadsheet detailing the group allocation for each animal can be sent directly to a third party nominated by the user, thus blinding the allocation. This enables the researcher to remain unaware of the group allocation until the data have been collected and analysed. (7) Diagrams can be safely shared with colleagues and collaborators at any stage of the process. (8) The user can export a PDF report, which contains key information about the internal validity of the experiment, a summary of the feedback from the system, and the EDA diagram itself. This report can be submitted as part of a grant application, as part of the ethical review process, or, later on, with a journal manuscript. Alternatively, the diagram data can be exported (as an.eda file) and saved locally or used to register the protocol before the experiment is conducted. (9) Once the planning is complete, the experiment is carried out. (10) The diagram can be updated after data collection to enable the user to keep an accurate record (e.g., to record the number of animals analysed if some failed to complete the experiment or if data are missing for other reasons).

<https://doi.org/10.1371/journal.pbio.2003779.g001>

publications. Novel ideas and intellectual property contained within the diagrams are protected; a summary of the security measures in place is included on the website: <https://eda.nc3rs.org.uk/security>. EDA diagrams are also computer interpretable, allowing automated critiquing of designs against recognised best practice without constraining the experimental design process or the plans themselves.

## The EDA enables an effective assessment of the experimental plans

For researchers who have limited access to statistical support, the critical feedback provided by the EDA will be particularly pertinent, as it provides users with information that is specific to the experiment they are planning. The system is not designed to replace specialist statistical advice but can facilitate it. The process of building a design using the EDA emulates the initial



**Fig 2. Example of an Experimental Design Assistant (EDA) diagram.** EDA diagrams are composed of nodes and links to represent an entire experimental plan. Each node contains properties where specific details are captured (properties are not shown in this picture, but in the EDA they are accessible by clicking on the specific node). This particular example is a simple 2-group comparison. The grey nodes contain high-level information about the experiment, such as the null and alternative hypotheses, the effect of interest (via the experiment node), the experimental unit, or the animal characteristics. The blue and purple nodes represent the practical steps carried out in the laboratory, such as the allocation to groups (allocation node) and the group sizes and role in the experiment (group nodes), the treatments (via the intervention nodes), and the measurements taken (measurement nodes). The green and red nodes represent the analysis, the outcome measures, and the independent variables.

<https://doi.org/10.1371/journal.pbio.2003779.g002>

fact-finding discussion a researcher might have with a statistician. It helps the researcher to identify much of the information that a statistician would need in order to provide expert advice and presents it in an explicit and standardised format, which can be made available to funding bodies, ethical review committees, journal editors, and peer reviewers.

Users also have the option to share their designs with team members and collaborators, and anecdotal evidence shows that diagrams are extremely useful when discussing the experimental plans within a research team. This is partly because the visual representation enables an efficient critical appraisal of the plans, such as questioning the role of and need for each experimental group, defining variables, identifying potential sources of bias, or debating the type of outcome measures. This detailed scrutiny before the experiment is carried out, or even before the plans are reviewed beyond the laboratory, enables researchers to identify potential pitfalls based on what they know about the science and experimental environment and perhaps follow up with more advanced questions to a statistician.

In addition to problems with the design, analysis, and reporting of scientific research, there are 2 practices that are widely encountered and further compromise the reliability of published results: 'p-hacking' (running multiple statistical tests on the same data and choosing the one



with the lowest  $p$  value) and selective outcome reporting (measuring different outcomes, or the same outcome in different ways, and only reporting the ones that reach statistical significance) [18]. These issues effectively represent a post hoc choice of outcome and analysis plan and would be prevented by formalising a clear protocol and plan for the statistical analysis before collecting the data. EDA diagrams are ideal for this purpose. EDA diagrams and the nonvisual information they contain (e.g., prespecified primary outcome measure, chosen method of randomisation) can be registered before the experiment is carried out, on specific platforms such as the Open Science Framework (<https://osf.io/>) or more generic platforms such as Figshare (<https://figshare.com>). This provides evidence that a study was planned as it has been reported and confirms that the primary outcome measure has not been changed during the course of the experiment and that any additional results reported should be treated with caution.

## The EDA promotes better understanding of experimental design and analysis

The EDA is not a ‘black box’ that instructs researchers on what design they should use. Instead, it promotes better understanding of experimental design and raises awareness about problems caused by a lack of randomisation and blinding, underpowered experiments, or inappropriate statistical analysis. The feedback provided by the system (see Table 2) enables users to learn about the implications of different design choices and helps them make informed decisions about the most appropriate one to adopt.

Animal studies often use suboptimal statistical analysis, such as failing to use factorial or block designs when appropriate, treating repeated measures as being independent, or failing to account for multiple testing [4,19–22]. The EDA encourages scientists to spend time planning their experiments and optimising the design interactively. It prompts researchers to evaluate carefully their experimental plan and to consider the data to be collected. It also helps to identify sources of variability by providing examples that are commonly encountered in animal research, such as the day of the experiment (if animals are used over several days); the time of day when the experiment is performed; the piece of equipment used to record measurements; the litter or cage mates; the location of cages in the room; or baseline variables, such as the animals’ weight or locomotor activity. Such sources of variability, termed ‘nuisance variables’, can then be accounted for in the design and analysis of the experiment (e.g., as covariates or blocking factors), or they can be standardised. The choice depends on the characteristics of the specific variable, for example, whether it can be treated as a continuous or categorical variable, and the extent of its likely impact on the variability of the response and on how far the conclusions of the experiment can be generalised. It also helps the user to identify independent variables that are repeated factors and warrant a repeated measure analysis, thereby ensuring that users are provided with enough information to avoid common pitfalls.

## Conclusion

The EDA is a novel tool bringing together machine-readable flow diagrams and computer-based logical reasoning to assist the robust and reproducible design of animal experiments. It ensures that the experimental plans are explicit and transparent, thus allowing greater scrutiny before and after data are collected and a meaningful dialogue between researchers and statisticians. It encourages improvements on the design by providing researchers with critical feedback and targeted information. Future development of the system will continue to incorporate user feedback to ensure that the EDA continues to support the needs of the research community. Together with comprehensive reporting and a better understanding of the factors that

**Table 2. Feedback provided by the Experimental Design Assistant (EDA).**

Aspect of experimental design covered by the feedback rules	High-level description of the type of feedback that the EDA can provide
Objective	Provides guidance to identify the null and alternative hypothesis, the effect of interest, and the effect size that is biologically relevant.
Randomisation	Detects when the allocation method is not specified. Highlights the importance of adequate randomisation, advises on randomisation procedures, and prompts users to consider different types of randomisation.
Blinding	Detects when there is no provision for blinding, explains why allocation concealment and blinding are important and the different stages of the experiment that can be blinded, and provides ways blinding can be achieved.
Groups and sample size	Provides guidance to identify the experimental unit(s) and determine suitable sample sizes.
Outcome measures	Highlights the implications of working with continuous and categorical data. Prompts user to identify the primary outcome measure.
Independent variables of interest	Detects when independent variables have not been identified or when they should be treated as continuous or categorical or as repeated factors. Detects when independent variables may be confounded.
Nuisance variables	Advises about nuisance variables commonly seen in animal experiments and how to account for them in the randomisation and analysis.
Statistical analysis	Suggests statistical analysis methods compatible with the design, along with software that can be used to perform these tests. Advises about parametric assumptions and data transformation. Suggests when the advice of a statistician should be sought.

The expert system provides helpful critical feedback to the user based on a set of rules. At the time of writing this article, approximately 140 feedback rules have been implemented; this number will increase over time, improving the precision of the feedback and the ability of the EDA to detect more subtle issues that could be addressed to optimise the design. The EDA feedback is split into 2 distinct functionalities: the Critique and the Analysis Suggestion.

The Critique helps the user build their diagram and identifies missing information and problems with internal consistency. It also suggests improvements to the design and assists the user in identifying and characterising the independent variables in the analysis. Feedback rules are devised to highlight the implications of different design choices, thus enabling researchers to make informed decisions. When an issue is detected (e.g., because additional information is requested or the information provided is not consistent with good practice), an error, warning, or advice flag is placed on a specific node in the diagram. Clicking on the flag opens a pop-up window explaining what the issue is and providing advice on how to address it. Not all EDA users will trigger all feedback rules; for instance, experienced researchers who initially input an adequate level of information within the diagram will see few, if any, prompts.

The Analysis Suggestion can be used after the Critique feedback has been addressed. The information provided is based on the number and type of independent variables of interest, nuisance variables, and outcome measures included in the analysis.

The development of the rule set was informed by workshops in which EDA diagrams representing examples of flawed experimental designs were analysed. These sessions identified, in particular, what information a statistician would require in order to offer the best advice on the design and what advice the statistician would feed back to the researcher to help them identify the information requested by the EDA or to help them improve the design. More information about the development of the feedback feature can be found here:

<https://eda.nc3rs.org.uk/feedback>

<https://doi.org/10.1371/journal.pbio.2003779.t002>



impact on the reliability and integrity of research findings, the EDA forms part of the solution identified by NC3Rs and others to improve the quality of animal research.

## Supporting information

**S1 Fig. Here, we present 4 Experimental Design Assistant (EDA) diagrams depicting ‘repeated measure designs’.** Diagram 1 shows an experiment in which animals receive multiple treatments in a different order for each animal, and each animal is used as its own control. Diagram 2 shows an experiment in which different groups of animals receive different treatments, 1 treatment per group, and the response to these treatments is measured over time, with each time point included in the analysis. Diagram 3 shows an experiment in which different groups of animals receive different treatments, 1 treatment per group, and the response to these treatments is measured over time, but a summary measure is taken for each animal. Diagram 4 shows an experiment in which multiple responses are measured for each animal. (PDF)

## References

1. Nature editorial. Announcement: Reducing our irreproducibility. *Nature*. 2013; 496(7446). Epub 2013/04/24.
2. Freedman LP, Cockburn IM, Simcoe TS. The economics of reproducibility in preclinical research. *PLoS Biol*. 2015; 13(6):e1002165. Epub 2015/06/10. <https://doi.org/10.1371/journal.pbio.1002165> PMID: 26057340; PubMed Central PMCID: PMC4461318.
3. Goodman SN, Fanelli D, Ioannidis JP. What does research reproducibility mean? *Science translational medicine*. 2016; 8(341):341ps12. Epub 2016/06/03. <https://doi.org/10.1126/scitranslmed.aaf5027> PMID: 27252173.
4. Kilkenny C, Parsons N, Kadyszewski E, Festing MF, Cuthill IC, Fry D, et al. Survey of the quality of experimental design, statistical analysis and reporting of research using animals. *PLoS ONE*. 2009; 4(11):e7824. Epub 2009/12/04. <https://doi.org/10.1371/journal.pone.0007824> PMID: 19956596.
5. Begley CG, Ellis LM. Drug development: Raise standards for preclinical cancer research. *Nature*. 2012; 483(7391):531–3. Epub 2012/03/31. <https://doi.org/10.1038/483531a> PMID: 22460880.
6. Begley CG, Ioannidis JP. Reproducibility in science: improving the standard for basic and preclinical research. *Circ Res*. 2015; 116(1):116–26. Epub 2015/01/02. <https://doi.org/10.1161/CIRCRESAHA.114.303819> PMID: 25552691.
7. NC3Rs. ARRIVE: Animal Research Reporting In Vivo Experiments 2017 [14 September 2017]. Available from: <https://www.nc3rs.org.uk/arrive-animal-research-reporting-vivo-experiments>.
8. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol*. 2010; 8(6):e1000412. Epub 2010/07/09. <https://doi.org/10.1371/journal.pbio.1000412> PMID: 20613859; PubMed Central PMCID: PMC2893951.
9. Cressey D. Surge in support for animal-research guidelines. *Nature* [Internet]. 2016. Available from: <http://www.nature.com/news/surge-in-support-for-animal-research-guidelines-1.19274>.
10. Cressey D. Web tool aims to reduce flaws in animal studies. *Nature* [Internet]. 2016; 531:[128 p.]. Available from: <http://www.nature.com/news/web-tool-aims-to-reduce-flaws-in-animal-studies-1.19459>.
11. Percie du Sert N, Bamsey I, Bate ST, Berdoy M, Clark RA, Cuthill IC, et al. The Experimental Design Assistant. *Nat Methods*. 2017; Epub Sept 28, 2017. <https://doi.org/10.1038/nmeth.4462>
12. Macleod MR, Lawson McLean A, Kyriakopoulou A, Serghiou S, de Wilde A, Sherratt N, et al. Risk of bias in reports of in vivo research: A focus for improvement. *PLoS Biol*. 2015; 13(10):e1002273. <https://doi.org/10.1371/journal.pbio.1002273> PMID: 26460723
13. Vesterinen HM, Sena ES, French-Constant C, Williams A, Chandran S, Macleod MR. Improving the translational hit of experimental treatments in multiple sclerosis. *Mult Scler*. 2010; 16(9):1044–55. Epub 2010/08/06. 1352458510379612 [pii] <https://doi.org/10.1177/1352458510379612> PMID: 20685763.
14. Rosenthal R, Fode KL. The effect of experimenter bias on the performance of the albino rat. *Behavioral Science*. 1963; 8(3):183–9. <https://doi.org/10.1002/bs.3830080302>

15. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci*. 2013; 14(5):365–76. Epub 2013/04/11. <https://doi.org/10.1038/nrn3475> PMID: 23571845.
16. Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005; 2(8):e124. Epub 2005/08/03. 04-PLME-E-0321R2 [pii] <https://doi.org/10.1371/journal.pmed.0020124> PMID: 16060722; PubMed Central PMCID: PMC1182327.
17. Bate ST, Clark RA. *The Design and Statistical Analysis of Animal Experiments*: Cambridge University Press; 2014.
18. John LK, Loewenstein G, Prelec D. Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychol Sci*. 2012; 23(5):524–32. Epub 2012/04/18. <https://doi.org/10.1177/0956797611430953> PMID: 22508865.
19. McCance I. Assessment of statistical procedures used in papers in the Australian Veterinary Journal. *Aust Vet J*. 1995; 72(9):322–8. Epub 1995/09/01. PMID: 8585846.
20. Festing MF. Randomized block experimental designs can increase the power and reproducibility of laboratory animal experiments. *ILAR J*. 2014; 55(3):472–6. Epub 2014/12/30. <https://doi.org/10.1093/ilar/ilu045> PMID: 25541548.
21. Shaw R, Festing MF, Peers I, Furlong L. Use of factorial designs to optimize animal experiments and reduce animal use. *ILAR J*. 2002; 43(4):223–32. Epub 2002/10/23. PMID: 12391398.
22. Nieuwenhuis S, Forstmann BU, Wagenmakers EJ. Erroneous analyses of interactions in neuroscience: a problem of significance. *Nat Neurosci*. 2011; 14(9):1105–7. Epub 2011/09/01. <https://doi.org/10.1038/nn.2886> PMID: 21878926.